Global Journal of Computer Science and Technology: B
Cloud and Distributed
Volume 16 Issue 3 Version 1.0 Year 2016
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

The Contemporary Review of Notable Cloud Resource Scheduling Strategies

By P.Sowjanya & Dr. K. V. N. Sunitha
University of Chittagong

Abstract- Cloud computing has become a revolutionary development that has changed the dynamics of business for the organizations and in IT infrastructure management. While in one dimension, it has improved the scope of access, reliability, performance and operational efficiency, in the other dimension, it has created a paradigm shift in the way IT systems are managed in an organizational environment. However, with the increasing demand for cloud based solutions, there is significant need for improving the operational efficiency of the systems and cloud based services that are offered to the customers. As cloud based solutions offer finite pool of virtualized on-demand resources, there is imperative need for the service providers to focus on effective and optimal resource scheduling systems that could support them in offering reliable and timely service, workload balancing, optimal power efficiency and performance excellence.

GJCST-B Classification: *J.5*

THECONTEMPORARYREVIEWOFNOTABLECLOUDRESOURCESCHEDULINGSTRATEGIES

Strictly as per the compliance and regulations of:

The Contemporary Review of Notable Cloud Resource Scheduling Strategies

P. Sowjanya α & Dr. K. V. N. Sunitha σ

Abstract- Cloud computing has become a revolutionary development that has changed the dynamics of business for the organizations and in IT infrastructure management. While in one dimension, it has improved the scope of access, reliability, performance and operational efficiency, in the other dimension, it has created a paradigm shift in the way IT systems are managed in an organizational environment. However, with the increasing demand for cloud based solutions, there is significant need for improving the operational efficiency of the systems and cloud based services that are offered to the customers. As cloud based solutions offer finite pool of virtualized on-demand resources, there is imperative need for the service providers to focus on effective and optimal resource scheduling systems that could support them in offering reliable and timely service, workload balancing, optimal power efficiency and performance excellence. There are numerous models of resource scheduling algorithms that has been proposed in the earlier studies, and in this study the focus is upon reviewing varied range of resource scheduling algorithms that could support in improving the process efficiency. In this manuscript, the focus is upon evaluating various methods that could be adapted in terms of improving the resource scheduling solutions.

I. Introduction

Information and cloud computing technology solutions has become an integral factor of today's communication systems and there are phenomenal developments in terms of how the companies are depending upon vivid range of information systems network for handling the business operations. Also, the volume of data that is generated by the companies is turning out to be a potential solution for the organizations. The volume of internet access and the big data solutions that are being implemented is creating a significant need for the organizations to focus upon varied kinds of data.

Such developments have eventually led to the rising demands of cloud computing solutions [1] [2]. The scope of computing resource capabilities and easy access to the data with increased mobility are the key solutions offered by cloud computing [3] and there are varied ways in which the cloud based solutions are offered to the users by the service providers. [4] [5].

Despite the fact that the cloud computing is similar to the process of grid computing and cluster computing models, in terms of characteristics that are common to parallel computing, but the usage of virtualization for resource management is a significant development [6] which facilitates in effective services as a utility model [7].

The scope of computing and accessibility to information has become much easier and the cost of managing IT systems to has come down significantly, which is not feasible in the traditional computing environment [2].

With the emerging trends of cloud computing solutions, there are numerous research studies that are being carried out in varied dimensions of cloud computing adaptation like focusing on increasing the operational scope of cloud computing, and predominantly the factor of virtualization. In the market oriented utility service stream like cloud computing, it is very important to focus upon optimally scheduling resources to reap potential benefits from the implementation [8].

In the SLA's between the service providers and end users, the emphasis is more on optimal scheduling of resources as the key deliverable of cloud computing services offered. [9]. Underestimating or unrealistic planning and provision of resources are leading to complexities [10]. In the dimension of optimum outcome, addressing the power efficient requirements, focusing on improving operational efficiency of the systems are becoming a major challenge for offering QoS requirements of the service. [11].

The process of improving the resource scheduling in cloud services has to focus upon identifying the suitable resources that are essential for scheduling an appropriate workload within time and also in terms of increasing the effective resource utilization process. In the other dimension, the quantum of resources availed for service offering also has to be minimal in terms of workload, ensuring required levels of service quality. To generate effective resource schedu- ling, best ways of mapping the resource workload is very important.

The second objective which plays a significant role in the resource scheduling process is to focus upon identifying the appropriate and suitable workload which can facilitate in scheduling multiple workloads which can address the QoS requirements like the CPU utilization, reliability, security and availability in the process of cloud services [12]. Hence it is very important that the resource scheduling focus upon the execution time for

Author α: Asst. Professor, Dept. of IT, GITAM University, Hyderabad. e-mail: sowjanya.ponnuru@gmail.com
Author σ: Professor and Principal, College of Engineering for Women, BVRIT, Hyderabad. e-mail: k.v.n.sunitha@gmail.com

every distinct workload and also predominantly for improving the overall performance which is depending upon the kind of workload like either the heterogeneous workloads or the homogenous workloads. [13].

There is extensive research and development in the process of cloud resource scheduling, and from the review of literature it is imperative that resource scheduling is one of the significant challenges facing the development of cloud based solutions. [14] [15]. The issue of Nondeterministic Polynomial Optimization Problem [16] [17] is considered to be one of the key challenges in the process of cloud resource scheduling, as NP-hard issues rise the usage by number of variables that are using the deterministic algorithm for exhaustive search. There are significant chances that there is dimensionality breakdown envisaged when the algorithms that solve relatively routine cloud scheduling problem. The issues are compounded when there is some kind of proliferation, ambition and complexity in terms of cloud computing.

Adapting an Evolutionary Computation Algorithm for tackling the cloud resource scheduling has grabbed the attention of researchers, as such a method can offer an effective solution to hard problems in the scheduling [18],[19],[20],[21],[22]. The efficacy of such systems in resource scheduling at grid computing levels [23], such algorithms have gained momentum for application in the cloud resource scheduling too [24] [25].

The surveys that are conducted on effective scheduling in cloud computing [26] [27] [28] [29] [26] and some of the methods like the "interconnected cloud computing" [30], not much of algorithms have been discussed in the process. In this study, the focus is on assessing the issues and systematic review of the existing models of cloud resource scheduling solutions that are being adapted in the cloud computing environment. The study focus on following key aspects.

- Taxonomic structure evaluation using the scheduling levels of hierarchy
- The clarity on the scope of cloud scheduling complications
- Survey of state-of-art approaches for handling cloud scheduling by evaluating the pros and cons
- Suggestions pertaining to how the various approach for varied levels of hierarchy of cloud resource scheduling is taking place.
- Analyzing the challenges pertaining to potential future research directions and real-time, adaptive dynamic, and distributed scheduling models.

II. Nomenclature of the Resource Scheduling in Cloud Computing

a) *Resource Scheduling*

In cloud resource scheduling among the key challenges that are envisaged in the process, dispersion, heterogeneity of resources and uncertainties are some of the key issues that are not addressed using the traditional resource scheduling algorithms (RSAs) [31] and it is imperative that there is need for making cloud oriented solutions that are more efficient in terms of caring the cloud environment properties. In the process of resource scheduling, the key stages are Resource Mapping, Resource Execution and Resource Monitoring.

Initially, the cloud consumer focus on the workload that is pending for execution, and followed by the process of mapping the workloads for right resources are carried out, using the QoS requirements agreed upon in the SLAs section of optimizing QoS parameters. Some of the important parameters like CPU utilization, throughput, memory utilization and other such factors are usually considered in the resource scheduling for all the cloud users in the network.

The prerogative of resource execution is to focus on allocating appropriate resources for suitable workloads on time, as the applications can focus on utilizing the resources effectively. Whilst executing a specific workload, the monitoring agent has to check the current workload. If the value of Required Resources (RRs) are higher than the value of Provided Resources (PRs) then it demands more resources. In such instances, the reserve resource pool maintained has to be used for providing the required resources using the rescheduling process, to achieve successful execution of the workload. Once the workload is completed, the resources that are free are released to the resource pool and the scheduler can focus on allocation the resources to execute new workloads. If efficient monitoring and utilization of computing resources are in place, it can help in improving the performance optimization. Thus it is imperative that there is need for effective and comprehensive intelligent monitoring agent for analyzing the performance of resource execution.

Profoundly, the SLAs should comprise information on varied deviations and scope defined from achieving the appropriate quality attributes. Cloud provider's SLA shall provide indication of the deviation of service if any, scope of feasibility for change and the factors to be considered whilst compensating any kind of unexpected outages [32]. To denote the CPU and Memory utilization, the resource monitoring agent is engaged, which collects the resource usage by evaluating performance metrics. It is very essential that the cloud provider have to focus on retaining adequate number of resources for delivering the continuous service to cloud consumer while addressing peak load.

Resource monitoring process shall be adapted for handling key QoS requirements whilst performing the workload execution. Two significant aspects of resource monitoring are about how the consumer prefers the execution of their workload with minimal cost and time, and without deviating from SLA's and the other in terms of how the provider plans to execute the workload with minimum resources.

Resource Monitoring is essential for analyzing performances both in terms of physical and virtual performance, as the resource utilization evaluation can lead to more effective ways of handling the processes. Also, the resource monitoring process can be adapted for handling varied factors like the security, reliability, approach and effectiveness, and confidentiality.

In the other way, the scope of achieving the process where the resources to be used for the process has to be minimum for a workload while adhering to the quality metrics of workload can be achieved only when the resource monitoring is effectively carried out. The process of deciding upon acquiring or releasing resources for workload, computing activities that are essential to be mapped for the cloud resources for improving the performance, as one of the key deliverables from the service provider is to focus on adhering to the compliance to SLA conditions pertaining to resource scheduling. [32].

From the above factors, it is imperative that there is need for effective resource scheduling algorithm that can support in managing the fluctuations for requirements in the workload and also towards maximizing utilization. To ensure that the resource scheduling requirements are effective, appropriate number of resources are to be deployed for executing the current load by addressing the challenges of under-loading or over-loading conditions..

b) *Hierarchy of Cloud Resources Scheduling*

In the process of resource management in a cloud computing environment, scheduling can be developed at varied levels of service stacks hierarchy. The architecture models of IaaS, PaaS, and SaaS stacks shall be adapted for classifying cloud scheduling problems in to the process of scheduling the application level, scheduling in virtualization layer and scheduling in the deployment layers. Table 1 indicates the hierarchy of cloud resources scheduling and the scope of scheduling process at varied hierarchy levels.

In terms of preceding the categories on the basis of high-level framework and taxonomy fro the cloud resource scheduling program, the low-level taxonomy can be achieved from varied range of scheduling objectives. Considering the implications like the deadlines and the budget constraints of the cloud users and also the resource needs that are to be balanced at a maximal rate by the service providers, the category of application layer resource scheduling is categorized in to further set of sub categories

- —“scheduling for user,”
- —“scheduling for provider efficiency,” and
- —“scheduling for negotiation.”

Table1: Nomenclature of the resource scheduling in cloud Computing

Nomenclature of the resource Scheduling	
Scheduling Approach	Objectives
Scheduling at Software layer	User QoS, Provider Efficiency, Negotiation
Scheduling at Platform Layer	Load Balance, Energy Efficiency Cost Effectiveness
Scheduling at Infrastructure Layer	Service Placement, Partner Federation, Data Routing

In the further process, at the scheduling levels in the virtualization layer, the challenges are predominantly about scheduling the Virtual Machines (VMs) and the Physical Machines (PMs) that has efficient load balance, level of conservation and the term of cost effectiveness. Hence the following sub categories are aimed in the process

- Scheduling for cost effectiveness
- Scheduling to ensure energy conservation
- Scheduling on basis of load balancing

And the deployment category of the scheduling process are sub categorized as

- Scheduling for partner federation
- Schedule for service placement
- Scheduling for data routing.

III. Review of Notable Cloud Resource Scheduling Strategies of Contemporary Solutions

Numerous scheduling algorithms are available for routine challenges in the cloud computing [34] [35] [36]. The algorithms that are offered in the models are little exhaustive in nature but can be very resourceful if the scheduling problem is manageable by converting to a combinational optimization problem like the Linear Programming [37] [38].

Integer Programming (IP) [39], and the Integer Linear Programming (ILP) [40] and the constraint satisfaction problem [41] are very effective models, but considering the NP-hard problem, cloud scheduling has to be addressed with enumerative approaches that can focus on increased dimensionality in terms of number of variables that are to be optimized.

a) *Bargaining Based Resource Scheduling*

Resource scheduling models has been developed by many of the earlier studies. Radu et.al [42] proposed CDA (Continuous Double Auction) model for distributed environment, which can support in executing scientific application where the negotiation between the place between resources and the scheduler, by focusing on self-limitation and aggressiveness.

Global Journal of Computer Science and Technology (B) Volume XVI Issue III Version I Year 2016

Scientific applications have the dependent tasks for which the output of one task is highly dependent on the other task. From the test results of implementing CDA model in CloudSim, the reduced time for completion and the reduced scope of relative error is imperative, but the model is more effective for homogenous workloads.

Lin et.al [43] developed a theoretical dynamic auction mechanism that can be very resourceful for handling the capacity distribution for evaluating the peak and off-peak demands depending on the capacity. Such mechanism supports in addressing the issues of computation capacity, but the model not focus on issues pertaining to any kind of deviation in SLA violation. Zhangjun et.al [44] has discussed market oriented based resource scheduling algorithm which contains the service and the task levels for dynamic resources for scheduling and assigning the task to service and the task to VM too. Such a method reduces the scope of operational costs for data centers and towards optimizing the makespan.

Mohsen et.al [45] has proposed market-oriented adaptive resource scheduling mechanisms for cost and time optimization in addition to addressing the deadlines for execution time. Such mechanism estimated the cost and time depending on completion time for different workloads on the basis of respective policies, but the process is limited only to one single IaaS provider that has uniform price.

Tdavid et.al [46] has proposed a model of distribution negotiation based resource scheduling model that focus on bargaining and attains higher utility. Such resource scheduling model can be resourceful for heterogeneous environment for improving the resource capabilities, cost and also time for completion. Seokho et.al [47] focus on the SLA oriented flexible negotiation that is based on resource scheduling pattern, and considers the crux of tradeoff among utilities for improving the speed and finding an effective service provider for quality performance service. Despite that the method has been resourceful in reducing SLA violations, still in terms of SLAs deviation there is rise in the kind of deviations.

b) Compromised Cost and Time based Resource Scheduling

The scheduling models that are proposed based on compromised cost and time has been proposed in some studies. Ganesh et.al [48] focused on pricing oriented scheduling algorithms, in which two self-evident bargaining methodologies has been discussed. Raiffa Bargaining Solution and Nash Bargaining solutions that are proposed in the study works on independent workflows.

Teng et.al [49] has worked an equilibrium resource scheduling technique for forecasting the prospect price even before knowing the competitors bidding information which has shown good results in the implementation on Cloudsim. HCOC (Hybrid Cloud Optimized Cost) model is proposed by Luiz et.al [50] which works on resource scheduling mechanism for addressing the problem of resource requirements that executes the workflows on the basis of budget and execution time, by focusing on adequate resource depending on QoS requirements.

In another model proposed by Ke et.al [51], the model emphasize on cost time based resource scheduling in which the cost constrained workflows are taken in to account, and execution time and cost for QoS parameters are considered.

c) Cost Based Resource Scheduling

Ana et. al [52] developed a model using the constraint resource algorithm, in which the First Come First Serve (FCFS) model has been adapted for reducing the cost, time required for completion and also for improving the CPU performance, still the challenge of starvation is a challenge in this model.

Ruben et.al [53] worked on optimization problem for imposing conditions like execution of job in a multi-provider hybrid cloud environment, depending on the requirements of data transmission, CPU and the memory, for categorization of non-provider and movable workloads.

Zhipiao et.al [54] has discussed an SLA aware genetic algorithm for resource scheduling mechanism that works on addressing the current requirement of varied applications based on virtual resources offered by third party infrastructure on lease model. Model has been effective in addressing the SLA violations and also in improving resource utilization and profit along with cost.

d) Dynamic and Adaptive Approaches forResource Scheduling

Ye et.al [57] developed a model of community aware resource scheduling technique with intension of reducing waiting time and average job slowdown time even without prior knowledge of real-time processing, for varied nodes being part of decentralized scheduling manner. Gaun et.al [58] has worked on queuing theory oriented model for improving the interval time average in an non interactive deadline-bound workload.

Altino et.al [59] has proposed failure and power aware resource scheduling model which works on reducing power consumption and adhering to SLAs. The levels of proactive fault-tolerance approaches used for decision making are effective in terms of handling failures for controlling the shared nodes.

Jiayin et.al [60] has proposed feedback based scheduling model for reducing any kind of resource contention issues using job preemption process. Ayasan et.al [61] has worked on Hadoop cluster based resource scheduling technique for calculating job arrival rate and also the execution time for making right decisions towards effective scheduling. The Hadoop system constituting a cluster, and is a combination of linked resources, is organized in to files and based on the file classification method for every job, the decision of whether a reduce task or map task is decided. Algorithm in the model focuses upon satisfaction of minimum share of requirements for all the users and fairness among the users in the system.

Zhen et.al [64] worked on virtualization oriented dynamic resource allocation mechanism for improving the server utilization. Skewness algorithm is adapted for estimating the disproportion in multi-dimensional utilization of a processor using hotspot mitigation. One of the key limitations in the model is about live migration related developments.

e) Energy Based Resource Scheduling

Energy based resource scheduling models are also profound solutions that has been worked in various models. Joseph et. al [65] has proposed a SLA aware machine learning based resource scheduling model for map reduction applications. In the proposed model, the exact solver depending on mixed linear programming aims to forecast the resource consumption based on current workload for executing varied tasks and responses time (taskSLA) for a workload and also taking in to account the contention among tasks that are executed on same resource.

Moreno et.al [66] focused upon the model of EASY (Energy Aware reconfiguration of software Systems) which is a QoS oriented resource scheduling technique towards reducing the power consumption. The model works on on-line algorithm for adjusting the processing speed of individual devices in dynamic manner to ensure that the average system response time is maintained well within the predefined threshold, while minimizing the total power consumption too

Yan et al. [67] in his study has proposed the model of controlled dependence graph which relies on the energy aware resource scheduling model for executing the HPC applications that are carried out with deadlines and also with scope of minimal energy consumption, in the distributed environment. Approximation of design and also the multiprocessor based scheduling algorithms are devised to address the problem based on the analysis and worst case performance assessment. Also, based on energy consumption, the desired deadline of tasks and pricing scheme is also designed for better execution.

In the other models [68] [69] [70] [71], there are numerous workload based resource scheduling policies that has been proposed for process improvement and also increasing the energy efficient methods adaptation. However, one of the focus areas in the models are about green revolution and improved performance.

f) Hybrid Approaches forResource Scheduling

Hybrid based resource scheduling solutions are developed considering varied metrics in to account. In many of the hybrid models that are proposed, the emphasis has been more on combination of data transfer, computational costs, and reduced cost factors. In the models that are proposed with Hybrid oriented approach, [72] [73] [74] the emphasis is on selection of resources from the public cloud for developing effective solutions on the basis of cloud based characteristics for task.

g) Heuristic and Meta-heuristic Approaches for Resource Scheduling

Heuristic based resource scheduling algorithms are very popular and in some of the models that are developed on heuristic methods, varied levels of heuristic methods were adapted for implementation [75] [76] [77] [78] for developing contemporary solutions using various kinds of algorithms. Distribution of resources to the workloads on the basis of requirements for reducing the execution time has been the focal point in the aforesaid models.

Raju et.al [79] proposed ACO and Cuckoo search for hybrid resource scheduling policy for reducing the completion time. Paulin et.al [80] proposed firefly oriented resource scheduling technique for improving the load balancing and also for execution time. Also, some of the parameters like load index, access rate, memory usage, processing time are some of the key factors that are taken in to consideration. Some of the other models like GA based [81] PSO based [82] [83] models discussed the scope of using specific algorithm models for reducing the execution time, and towards improving user's satisfaction.

In [84] [85] [86] [87], also, there were numerous models of GA solutions that are proposed for handling the independent and divisible tasks for cloud computing environment. The models assert the fact that GA approach could be resourceful in handling the user costs together, and also in terms of improving the resourcefulness in a cloud computing environment.

Also, in [88] [89], the models pertaining to adapting the cloud resources with GA-based approach by focusing on virtual resources for scheduling has been proposed with impressive test outcomes. In [90] the model focus on using VM match and execution orders for resource scheduling.

A general framework of using ACO to schedule user tasks is as follows, each ant uses M steps to construct a solution. In the i^{th} step to schedule the i^{th} task T_i, the ant uses pheromone and heuristic information to choose the suitable resource R_j. After M steps, all the M tasks have been scheduled on different resources.

Taking stock of above factors, in [91] [92] there are many scheduled M tasks one by one for the cloud resources, using the scheme as at every step for the task can be scheduled on the basis of resources that are set. The heuristic information based on users QoS metrics for user cost, system reliability, response and security are adapted for guiding the ant for selecting optimal resource which is proposed in [92] [93]. The tasks were classified in to varied categories are bound to be adapted on cloud resources using ACO optimization.

In terms of load balancing for resource scheduling in [103], Nishant et.al proposes an optimized load balancing method using the ACO based algorithm. In extensive of the aforesaid model, in [104] the model has been extended with another objective of provider efficiency oriented scheduling of resource utilization.

Wen et.al [105] has focused on improving the cloud resource utilization ratio by scheduling the cloud resources on the basis of hybrid algorithm comprising ACO and PSO. One of the key objectives of developing provider efficiency for oriented scheduling is to work on energy consumption for the cloud center [106]. Such tasks are usually scheduled using modified GA along with local search for optimizing the energy consumption and for reaping potential benefits from the outcome in terms saving energy, by using the combination of ACO and Cuckoo search. [107]. The ACO approach that is adapted as main framework and the application of CS rather than heuristic information for finding the next resource for task is an effective outcome.

h) Observations

From the review of the studies, it is evident that despite of numerous models that are proposed, majority of them are probabilistic rather than deterministic. Also not every scope and parameter of QoS metrics in SLAs are assured in the existing models. In this research review, the focus is upon reviewing the earlier models in the area of cloud resource scheduling and the QoS metrics adapted in the process of effective resource management. It is also evident that by using some of the effective evolutionary models of resource scheduling algorithms, the operational efficiency of resource scheduling can be improved to great extent.

IV. Conclusion

Information systems management has become an integral part of organizational process and with the kind of fast emerging computational solutions and methods like the cloud computing solutions, the efficacy of the systems has increased standards of accessibility, flexibility and scalability. In the contemporary scenario, organizations are keen on adapting cloud based solutions for the information systems management in the context of various reasons like the ease of access to data, reduced cost of operations. And IT infrastructure management related costs going down, and the flexibility in terms of managing the data in third party services. However, one of the significant factors that make a vital role in the success of adapting the cloud computing process is about the Service Level Agreements with the service provider for offering the reliable services.

Despite the fact that there are many factors that impact the service quality, one of the profound factors that impact the cloud solution efficacy is about resource scheduling process in the cloud services. There are numerous models that have been proposed in the earlier models. It is imperative from the review of extensive literature on the varied models of cloud based resource scheduling algorithms, that in the case of majority of models that are proposed, they are probabilistic models rather than being a deterministic models.

Considering the fact that the current models have high levels of computational complexities and the rising standards and requirements of cloud based solutions, there is significant need for focusing on more effective solutions that could be adapted for resource scheduling. Using some of the effective methods like CUCKOO search, TABU search and other such contemporary evolutionary algorithms, the process of resource scheduling can be improvised to great extent. If such a contemporary solution could be achieved it can support in improving the efficiency and attaining the optimal resource scheduling process in the cloud services and also in terms of reducing the linear complexity in the systems.

Références

1. Dragland. 2013. Big data, for better or worse: 90% of world's data generated over last two years. Science Daily (May 2013).
2. N. Toosi, R. N. Calheiros, and R. Buyya. 2014. Interconnected cloud computing environments: Challenges, taxonomy, and survey. ACM Computing Surveys 47, 1 (2014), 1–47.
3. Z. H. Zhang and X. J. Zhang. 2010. A load balancing mechanism based on ant colony and complex network theory in open cloud computing federation. In Proceedings of the 2nd International Conference on Industrial Mechatronics and Automation. 240–243.
4. Foster, Y. Zhao, I. Raicu, and S. Lu. 2008. Cloud computing and grid computing 360-degree compared. In Proceedings of the Grid Computing Environments Workshop, 1–10.
5. F. Xu, F. M. Liu, H. Jin, and A. V. Vasilakos. 2014. Managing performance overhead of virtual machines in cloud computing: A survey, state of the art, and future directions. Proceedings of the IEEE 102, 1 (2014), 11–31.
6. T. Grandison, E. M. Maximilien, S. Thorpe, and A. Alba. 2010. Towards a formal definition of a computing cloud. In Proceedings of the IEEE 6th World Congress on Services. 191–192.
7. R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandicc. 2009. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering

computing as the 5th utility. Future Generation Computer Systems 25 (2009), 599–616.
8. M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, et al. 2010. A view of cloud computing. Communications of the ACM 53, 4 (2010), 50–58.
9. H. Morshedlouand M. R.Meybodi. 2014. Decreasing impact of SLA violations: A proactive resource allocation approach for cloud computing environments. IEEE Transactions on Cloud Computing 2, 2 (2014), 156– 167.
10. M. D. Dikaiakos, G. Pallis, D. Katsaros, P. Mehra, and A. Vakali. 2009. Cloud computing: Distributed Internet computing for IT and scientific research. IEEE Internet Computing 13, 5 (2009), 10–13.
11. J. Baliga, R. W. A. Ayre, K. Hinton, and R. S. Tucker. 2011. Green cloud computing: Balancing energy in processing, storage, and transport. Proceedings of the IEEE 99, 1 (2011), 149–167.
12. Singh, S., Chana, I.: QoS-aware autonomic resource management in cloud computing: a systematic review. ACM Comput. Surv. 48(3), 39 (2015)
13. Singh, S., Chana, I.: Cloud resource provisioning: survey, status and future research directions. Knowl. Inf. Syst. 44, 1–50 (2015)
14. L. Heilig and S. Vob. 2014. A scientometric analysis of cloud computing literature. IEEE Transactions on Cloud Computing 2, 3 (2014), 266–278.
15. N. Toosi, R. N. Calheiros, and R. Buyya. 2014. Interconnected cloud computing environments: Challenges, taxonomy, and survey. ACM Computing Surveys 47, 1 (2014), 1–47.
16. T. A. L. Genez, L. F. Bittencourt, and E. R. M. Madeira. 2012. Workflow scheduling for SaaS/PaaS cloud providers considering two SLA levels. In Proceedings of the IEEE Network Operations and Management Symposium. 906–912.
17. J. L. Xu, J. Tang, K. Kwiat, W. Y. Zhang, and G. L. Xue. 2013. Enhancing survivability in virtualized data centers: A service-aware approach. IEEE Journal on Selected Areas in Communications 31, 12 (2013), 2610–2619.
18. T. Back, M. Emmerich, and O. M. Shir. 2008. Evolutionary algorithms for real world applications. IEEE Computational Intelligence Magazine 3, 1 (2008), 64–67.
19. J. Zhang, Z. H. Zhan, Y. Lin, N. Chen, Y. J. Gong, J. H. Zhong, H. Chung, Y. Li, and Y. H. Shi. 2011c. Evolutionary computation meets machine learning: A survey. IEEE Computational Intelligence Magazine 6, 4 (2011), 68–75.
20. Y. L. Li, Z. H. Zhan, Y. J. Gong, W. N. Chen, J. Zhang, and Y. Li. 2014. Differential evolution with an evolution path: A DEEP evolutionary algorithm. IEEE Transactions on Cybernetics, DOI: 10.1109/TCYB.2014.2360752
21. N. Chen, W. N. Chen, Y. J. Gong, Z. H. Zhan, J. Zhang, Y. Li, and Y. S. Tan. 2014. An evolutionary algorithm with double-level archives for multiobjective optimization. IEEE Transactions on Cybernetics, DOI: 10.1109/TCYB.2014.2360923
22. Y. H. Li, Z. H. Zhan, S. J. Lin, J. Zhang, and X. N. Luo. 2015c. Competitive and cooperative particle swarm optimization with information sharing mechanism for global optimization problems. Information Sciences 293 (2015), 370–382.
23. J. Yu, R. Buyya, and K. Ramamohanarao. 2008. Workflow scheduling algorithms for grid computing. Metaheuristics for Scheduling in Distributed Computing Environments Studies in Computational Intelligence 146 (2008), 173–214.
24. Jennings and R. Stadler. 2014. Resource management in clouds: Survey and research challenges. Journal of Network Systems Management 1–53.
25. M. A. Rodriguez and R. Buyya. 2014. Deadline based resource provisioning and scheduling algorithm for scientific workflows on clouds. IEEE Transactions on Cloud Computing 2, 2 (2014), 222–235.
26. J. Baliga, R. W. A. Ayre, K. Hinton, and R. S. Tucker. 2011. Green cloud computing: Balancing energy in processing, storage, and transport. Proceedings of the IEEE 99, 1 (2011), 149–167.
27. K. Bardsiri and S. M. Hashemi. 2012. A review of workflow scheduling in cloud computing environment. International Journal of Computer Science and Management Research 1, 3 (2012), 348–351.
28. H. Kaur and M. Singh. 2012. Review of various scheduling techniques in cloud computing. International Journal of Networking & Parallel Computing 1, 2 (2012).
29. Y. Chawla and M. Bhonsle. 2012. A study on scheduling methods in cloud computing. International Journal of Emerging Trends & Technology in Computer Science 1, 3 (2012), 12–17.
30. F. Xu, F. M. Liu, H. Jin, and A. V. Vasilakos. 2014. Managing performance overhead of virtual machines in cloud computing: A survey, state of the art, and future directions. Proceedings of the IEEE 102, 1 (2014), 11–31.
31. Singh, S., Chana, I.: QRSF: QoS-aware resource scheduling framework in cloud computing. J. Supercomput. 71(1), 241–292 (2015)
32. Singh, S., Chana, I.: Cloud resource provisioning: survey, status and future research directions. Knowl. Inf. Syst. 44, 1–50 (2015)
33. N. Toosi, R. N. Calheiros, and R. Buyya. 2014. Interconnected cloud computing environments: Challenges, taxonomy, and survey. ACM Computing Surveys 47, 1 (2014), 1–47.
34. K. Bardsiri and S. M. Hashemi. 2012. A review of workflow scheduling in cloud computing environment. International Journal of Computer Science and Management Research 1, 3 (2012), 348–351.
35. H. Kaur and M. Singh. 2012. Review of various scheduling techniques in cloud computing. Interna-

tional Journal of Networking & Parallel Computing 1, 2 (2012).
36. Y. Chawla and M. Bhonsle. 2012. A study on scheduling methods in cloud computing. International Journal of Emerging Trends & Technology in Computer Science 1, 3 (2012), 12–17.
37. S. Kumar and P. Balasubramanie. 2012. Dynamic scheduling for cloud reliability using transportation problem. Journal of Computer Science 8, 10 (2012), 1615–1626.
38. Speitkamp and M. Bichler. 2010. A mathematical programming approach server consolidation problems in virtualized data centers. IEEE Transactions on Services Computing 3, 4 (2010), 266–278.
39. Q. Li and Y. K. Guo. 2010. Optimization of resource scheduling in cloud computing. In Proceedings of the 12th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing. 315–320.
40. T. A. L. Genez, L. F. Bittencourt, and E. R. M. Madeira. 2012. Workflow scheduling for SaaS/PaaS cloud providers considering two SLA levels. In Proceedings of the IEEE Network Operations and Management Symposium. 906–912.
41. H. N. Van, F. D. Tran, and J. Menaud. 2010. Performance and power management for cloud infrastructures. In Proceedings of the IEEE 3rd International Conference on Cloud Computing. 329–336.
42. Prodan, R., Wieczorek, M., Fard, H.M.: Double auction based scheduling of scientific applications in distributed grid and cloud environments. J. Grid Comput. 9(4), 531– 548 (2011)
43. Lin, W.-Y., Lin, G.-Y., Wei, H.-Y.: Dynamic auction mechanism for cloud resource allocation. In: 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing (CCGrid), pp. 591–592. IEEE (2010)
44. Wu, Z., Liu, X., Ni, Z., Yuan, D., Yang, Y.: A market-oriented hierarchical scheduling strategy in cloud workflow systems. J. Supercomput. 63 (1), 256–293 (2013)
45. Salehi, M.A., Buyya, R.: Adapting market-oriented scheduling policies for cloud computing. In: Algorithms and Architectures for Parallel Processing, pp. 351–362. Springer, Berlin (2010)
46. An, B., Lesser, V., Irwin, D., Zink, M.: Automated negotiation with decommitment for dynamic resource allocation in cloud computing. In: Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1, vol. 1, pp. 981–988. International Foundation for Autonomous Agents and Multiagent Systems (2010)
47. Son, S., Jun, S.C.: Negotiation-based flexible SLA establishment with SLA-driven resource allocation in cloud computing. In: 2013 13th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid), pp. 168–171. IEEE (2013)
48. Iyer, G.N., Veeravalli, B.: On the resource allocation and pricing strategies in Compute Clouds using bargaining approaches. In: 2011 17th IEEE International Conference on Networks (ICON), pp. 147–152. IEEE (2011)
49. Teng, F., Magoules, F.: Resource pricing and equilibrium allocation policy in cloud computing. In: 2010 IEEE 10th International Conference on Computer and Information Technology (CIT), pp. 195–202. IEEE (2010)
50. Bittencourt, L.F., Madeira, E.R.M.: HCOC: a cost optimization algorithm for workflow scheduling in hybrid clouds. J. Internet Serv. Appl. 2(3), 207–227 (2011)
51. Liu, K., Jin, H., Chen, J., Liu, X., Yuan, D., Yang, Y.: A compromised-time-cost scheduling algorithm in SwinDeW-C for instance-intensive cost-constrained
52. Oprescu, A., Kielmann, T.: Bag-of-tasks scheduling under budget constraints. In: 2010 IEEE Second International Conference on Cloud Computing Technology and Science (CloudCom), pp. 351–359. IEEE (2010)
53. Van den Bossche, R., Vanmechelen, K., Broeckhove, J.: Cost-optimal scheduling in hybrid iaas clouds for deadline constrained workloads. In: 2010 IEEE 3rd International Conference on Cloud Computing (CLOUD), pp. 228–235. IEEE (2010)
54. Liu, Z., Wang, S., Sun, Q., Zou, H., Yang, F.: Cost-aware cloud service request scheduling for SaaS providers. Comput. J. 57(2), bxt009 (2013)
55. Su, S., Li, J., Huang, Q., Huang, X., Shuang, K., Wang, J.: Cost-efficient task scheduling for executing large programs in the cloud. Parallel Comput. 39(4), 177–188 (2013)
56. Moschakis, I.A., Karatza, H.D.: Performance and cost evaluation of Gang Scheduling in a Cloud Computing system with job migrations and starvation handling. In: 2011 IEEE Symposium on Computers and Communications (ISCC), pp. 418–423. IEEE (2011)
57. Huang, Y., Bessis, N., Norrington, P., Kuonen, P., Hirsbrunner, B.: Exploring decentralized dynamic scheduling for grids and clouds using the community-aware scheduling algorithm. Futur. Gener. Comput. Syst. 29(1), 402–415 (2013)
58. Le, G., Xu, K., Song, J.: Dynamic resource provisioning and scheduling with deadline constraint in elastic cloud. In: 2013 International Conference on Service Sciences (ICSS), pp. 113–117. IEEE (2013)
59. Sampaio, A.M., Barbosa, J.G.: Dynamic power-and failure-aware cloud resources allocation for sets of independent tasks. In: 2013 IEEE International Conference on Cloud Engineering (IC2E), pp. 1–10. IEEE (2013)
60. Li, J., Qiu, M., Niu, J., Gao, W., Zong, Z., Qin, X.: Feedback dynamic algorithms for preemptable job scheduling in cloud systems. In: 2010 IEEE/WIC/

ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), vol. 1, pp. 561–564. IEEE (2010)
61. Rasooli, A., Down, D.: An adaptive scheduling algorithm for dynamic heterogeneous Hadoop systems. In: Proceedings of the 2011 Conference of the Center for Advanced Studies on Collaborative Research, pp. 30–44. IBM Corp (2011)
62. Lee, Z., Wang, Y., Zhou, W.: A dynamic priority scheduling algorithm on service request scheduling in cloud computing. In: 2011 International Conference on Electronic and Mechanical Engineering and Information Technology (EMEIT), vol. 9, pp. 4665–4669. IEEE (2011)
63. Hwang, J., Wood, T.: Adaptive dynamic priority scheduling for virtual desktop infrastructures. In: Proceedings of the 2012 IEEE 20th International Workshop on Quality of Service, p. 16. IEEE Press (2012)
64. Xiao, Z., Song, W., Qi, C.: Dynamic resource allocation using virtual machines for cloud computing environment. IEEE Trans. Parallel Distrib. Syst. 24(6), 1107– 1117 (2013)
65. Rahman, M., Hassan, R., Ranjan, R., Buyya, R.: Adaptive workflow scheduling for dynamic grid and cloud computing environment. Concurr. Comput.: Pract. Exp. 25(13), 1816–1842 (2013)
66. Marzolla, M., Mirandola, R.: Dynamic power management for QoS-aware applications. Sustain. Comput.: Inf. Syst. 3(4), 231–248 (2013)
67. Ma, Y., Gong, B., Sugihara, R., Gupta, R.: Energyefficient deadline scheduling for heterogeneous systems. J. Parallel Distrib. Comput. 72(12), 1725–1740 (2012)
68. Ying, C., Jiong, Y.: Energy-aware genetic algorithms for task scheduling in cloud computing. In: 2012 Seventh ChinaGrid Annual Conference (ChinaGrid), pp. 43–48. IEEE (2012)
69. Kim, N., Cho, J., Seo, E.: Energy-credit scheduler: an energy-aware virtual machine scheduler for cloud systems. Futur. Gener. Comput. Syst. 32, 128–137 (2014)
70. Yassa, S., Chelouah, R., Kadima, H., Granado, B.: Multi-objective approach for energy-aware workflow scheduling in cloud computing environments. Sci. World J. 2013(Article ID 350934), 13 (2013). doi:10. 1155/2013/350934
71. Chen, C., He, B., Tang, X.: Green-aware workload scheduling in geographically distributed data centers. In: CloudCom, pp. 82–89 (2012). doi:10. 1155/2013/350934
72. Van den Bossche, R., Vanmechelen, K., Broeckhove, J.: Cost-efficient scheduling heuristics for deadline constrained workloads on hybrid clouds. In: Cloud Computing Technology and Science (CloudCom), 2011 IEEE Third International Conference on, pp. 320–327. IEEE (2011)
73. Calheiros, R.N., Buyya, R.: Cost-effective provisioning and scheduling of deadline-constrained applications in hybrid clouds. In:Web Information Systems Engineering- WISE 2012, pp. 171–184. Springer, Berlin (2012)
74. Kumar, B.A., Ravichandran, T.: Time and cost optimization algorithm for scheduling multiple workflows in hybrid clouds. Eur. J. Sci. Res. 89(2), 265–275 (2012)
75. Xu, G., Ding, Y., Zhao, J., Hu, L., Fu, X.: A novel artificial bee colony approach of live virtual machine migration policy using Bayes theorem. Sci.World J. 2013(Article ID 369209), 13 (2013). doi:10.1155/ 2013/369209
76. Song, X., Gao, L., Wang, J.: Job scheduling based on ant colony optimization in cloud computing. In: 2011 International Conference on Computer Science and Service System (CSSS), pp. 3309–3312. IEEE (2011)
77. Nishant, K., Sharma, P., Krishna, V., Gupta, C., Singh, K.P., Rastogi, R.: Load balancing of nodes in cloud using ant colony optimization. In: 2012 UKSim 14th International Conference on Computer Modelling and Simulation (UKSim), pp. 3–8. IEEE (2012)
78. Bitam, S.: Bees Life Algorithm for job scheduling in cloud computing. In: International Conference on Computing and Information Technology. ICCIT, pp 186–191 (2012)
79. Raju, R., Babukarthik, R.G., Chandramohan, D., Dhavachelvan, P., Vengattaraman, T.: Minimizing the makespan using Hybrid algorithm for cloud computing. In: 2013 IEEE 3rd International Advance Computing Conference (IACC), pp. 957–962. IEEE (2013)
80. Szabo, C., Sheng, Q.Z., Kroeger, T., Zhang, Y., Jian, Y.: Science in the cloud: allocation and execution of dataintensive scientific workflows. J. Grid Comput. 12(2), 245–264 (2014)
81. Morariu, O., Morariu, C., Borangiu, T.: A genetic algorithm for workload scheduling in cloud based e-learning. In: Proceedings of the 2nd International Workshop on Cloud Computing Platforms, p. 5. ACM (2012)
82. Somasundaram, T.S., Govindarajan, K.: CLOUDRB: a framework for scheduling and managing High-Performance Computing (HPC) applications in science cloud. Futur. Gener. Comput. Syst. 34, 47–65 (2014)
83. Netjinda, N., Sirinaovakul, B., Achalakul, T.: Cost optimal scheduling in IaaS for dependent workload with particle swarm optimization. J. Supercomput. 68(3), 1579–1603 (2014)
84. H. Zhao, S. S. Zhang, Q. F. Liu, J. Xie, and J. C. Hu. 2009. Independent tasks scheduling based on genetic algorithm in cloud computing. In Proceedings of the 5th International Conference on Wire-

less Communications, Networking and Mobile Computing, 1–4.
85. P. Kumar and A. Verma. 2012. Independent task scheduling in cloud computing by improved genetic algorithm. International Journal of Advanced Research in Computer Science and Software Engineering 2, 5 (2012), 111–114.
86. G.N. Gan, T. L. Huang, and S. Gao. 2010. Genetic simulated annealing algorithm for task scheduling based on cloud computing environment. In Proceedings of the International Conference on Intelligent Computing and Integrated Systems. 60–63.
87. J.W. Ge and Y. S. Yuan. 2013. Research of cloud computing task scheduling algorithm based on improved genetic algorithm. In Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering. 2134–2137.
88. Barrett, E. Howley, and J. Duggan. 2011. A learning architecture for scheduling workflow applications in the cloud. In Proceedings of the 9th IEEE European Conference on Web Services. 83–90.
89. Z. Ye, X. F. Zhou, and A. Bouguettaya. 2011. Genetic algorithm based QoS-Aware service compositions in cloud computing. Database Systems for Advanced Applications, Lecture Notes in Computer Science, Volume 6588. Springer, Berlin, 321–334.
90. Szabo and T. Kroeger. 2012. Evolving multi-objective strategies for task allocation of scientific workflows on public clouds. In Proceedings of the IEEE World Congress on Computation Intelligence. 1–8.
91. S. Banerjee, I. Mukherjee, and P. K. Mahanti. 2009. Cloud computing initiative using modified ant colony framework. World Academy of Science, Engineering and Technology 56 (2009), 221–224.
92. H. Liu, D. Xu, and H. K. Miao. 2011. Ant colony optimization based service flow scheduling with various QoS requirements in cloud computing. In Proceedings of the 1st ACIS International Symposium on Software and Network Engineering. 53–58.
93. L. N. Zhu, Q. S. Li, and L. N. He. 2012. Study on cloud computing resource scheduling strategy based on the ant colony optimization algorithm. International Journal of Computer Science Issues 9, 5 (2012), 54–58.
94. S. Pandey, L. L. Wu, S. M. Guru, and R. Buyya. 2010. A particle swarm optimization-based heuristic for scheduling workflow applications in cloud computing environments. In Proceedings of the 24th IEEE International Conference on Advanced Information Networks and Applications. 400–407.
95. Z. J. Wu, Z. W. Ni, L. C. Gu, and X. Liu. 2010. A revised discrete particle swarm optimization for cloud workflow scheduling. In Proceedings of the International Conference on Computational Intelligence and Security. 184–188.
96. W. N. Chen and J. Zhang. 2012. A set-based discrete PSO for cloud workflow scheduling with user-defined QoS constraints. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics. 773–778.
97. L. Z. Guo, S. G. P. Zhao, S. G. Shen, and C. Y. Jiang. 2012. A particle swarm optimization for data placement strategy in cloud computing. Information Engineering and Applications, Lecture Notes in Electrical Engineering, 323–330.
98. M. A. Rodriguez and R. Buyya. 2014. Deadline based resource provisioning and scheduling algorithm for scientific workflows on clouds. IEEE Transactions on Cloud Computing 2, 2 (2014), 222–235.
99. H. H. Li, Y. W. Fu, Z. H. Zhan, and J. J. Li. 2015a. Renumber strategy enhanced particle swarm optimization for cloud computing resource scheduling. In Proceedings of the IEEE Congress on Evolution Computation, in press.
100. H. H. Li, Z. G. Chen, Z. H. Zhan, K. J. Du, and J. Zhang. 2015b. Renumber coevolutionarymultiswarm particle swarm optimization for multi-objective workflow scheduling on cloud computing environment. In Proceedings of the Genetic Evolutionary Computation Conference.
101. Z. H. Zhan, J. Li, J. Cao, J. Zhang, H. Chung, and Y. H. Shi. 2013. Multiple populations for multiple objectives: A coevolutionary technique for solving multiobjective optimization problems. IEEE Transactions on Cybernetics 43, 2 (2013), 445–463.
102. K. Zhu, H. G. Song, L. J. Liu, J. Z. Gao, and G. J. Cheng. 2011. Hybrid genetic algorithm for cloud computing applications. In Proceedings of the IEEE Asia-Pacific Services Computing Conference. 182–187.
103. K. Nishant, P. Sharma, V. Krishna, C. Gupta, K. P. Singh, N. Nitin, and R. Rastogi. 2012. Load balancing of nodes in cloud using ant colony optimization. In Proceedings of the 14th International Conference on Computer Modelling and Simulation. 3–8.
104. Q. C. Lv, X. X. Shi, and L. Z. Zhou. 2012. Based on ant colony algorithm for cloud management platform resources scheduling. In Proceedings of the World Automation Congress. 1–4.
105. X. T. Wen, M. H. Huang, and J. H. Shi. 2012. Study on resources scheduling based on ACO algorithm and PSO algorithm in cloud computing. In Proceedings of the 11th International Symposium on Distributed Computing and Applications to Business, Engineering & Science. 219–222.
106. G. Shen and Y. Q. Zhang. 2011. A shadow price guided genetic algorithm for energy aware task scheduling on cloud computers. In Proceedings of the International Conference on Advances in Swarm Intelligence. Lecture Notes in Computer Science, Volume 6728. Springer, Berlin, 522–529.
107. X. L. Wang, Y. P. Wang, and H. Zhu. 2012. Energy-efficient multi-job scheduling model for cloud com-

Global Journal of Computer Science and Technology (B) Volume XVI Issue III Version I Year 2016

puting and its genetic algorithm. Mathematical Problems in Engineering, Article ID 589243, 1–16.
108. R. G. Babukarthik, R. Raju, and P. Dhavachelvan. 2012. Energy-aware scheduling using hybrid algorithm for cloud computing. In Proceedings of the 3rd International Conference on Computing Communication & Networking Technologies. 1–6.

This page is intentionally left blank